J. Phys. Chem. A **1999,** *103,* 103−118

The Role of Bonded Terms in Free Energy Simulations: 1. Theoretical Analysis

Stefan Boresch[†,‡,§] **and Martin Karplus***[,†,‡]

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138, Laboratoire de Chimie Biophysique, Institut Le Bel, Université Louis Pasteur, 67000 Strasbourg, France, and Institut für Theoretische Chemie, University of Vienna, Währingerstrasse 17, A-1090 Vienna, Austria

Received: March 25, 1998

The role of the bonded (bond stretching and bond angle) force-field terms in free energy simulations is examined. It is shown that the proper treatment of such terms depends on the choice of the free energy methodology (single or dual topology). Furthermore, while there are no problems in describing changes in bonded terms, care has to be used in creating or destroying them in a molecular dynamics simulation. An approach that avoids the singularity caused by a bond with a zero force constant is outlined. Changes in bond stretching or bond angle terms are shown to give rise to vibrational, Jacobian factor, and potential-of-mean-force-type (pmf) contributions. The meaning of bond stretching and bond angle bending free energy components obtained in single and dual topology simulations and their connection to these three contributions is investigated. Due to the different end states used in single and dual topology simulations, the pmf contribution is projected on different free energy components. In certain dual topology methods, vibrational and Jacobian factor contributions are not included in the free energy difference. Therefore, single free energy differences (e.g., the free energy difference between two molecules in the gas phase and in solution) often cannot be compared directly between single and dual topology methods. However, identical double free energy differences (e.g., free energy differences of solvation) are obtained in all cases. The present analysis emphasizes the importance of the details of the simulation methodology in interpreting the results for bonded terms and reconciles apparently contradictory findings in the literature.

1. Introduction

Free energy simulations have been successfully applied to calculate the free energy change for a variety of processes.[1−4] They include protein stability,[5−7] ligand binding,[8,9] cooperativity,[10] solvation,[11] and conformational equilibria.[12] Such simulations make it possible to probe the systems of interest at the atomic level of detail and to obtain insights that would be difficult to obtain experimentally. Although most calculated overall free energy changes have been made in response to experimental measurements,[4−12] examples of predictions of free energy differences have been published (e.g., the free energy difference of solvation between nucleic acid bases[13,14] or a study of the thermodynamics of Ribonuclease T1 substrate interactions).[15] Free energy simulations also have been used in force-field development.[16,17]

Most free energy simulations are concerned with *alchemical* transformations,[3,10] in which one calculates the free energy change resulting from the transformation of one molecule or one molecular fragment (e.g., an amino acid side chain) into another. The physically significant quantity (e.g., the free energy difference of solvation between two solutes or the free energy difference of binding for two ligands) can be expressed as a double free energy difference in almost all cases. For example, the free energy difference of solvation, $\Delta\Delta A_{\mathrm{solv}}$, between two solutes S1 and S2 is defined as

$$\Delta\Delta A_{\mathrm{solv}} = \Delta A_2 - \Delta A_1 \tag{1}$$

where ΔA_1 and ΔA_2 are the free energies of transfer S1 and S2, respectively, from the gas phase into aqueous solution. Since free energies are state functions, a thermodynamic cycle (see Figure 1) can be used to obtain $\Delta\Delta A_{\mathrm{solv}}$ from the alchemical free energy differences calculated as the difference between the free energy change of mutating S1 into S2 in solution (ΔA_4) and in the gas phase (ΔA_3);[18] i.e.,

$$\Delta\Delta A_{\mathrm{solv}} = \Delta A_4 - \Delta A_3 \tag{2}$$

Similarly, the free energy difference for the binding of two ligands, L1 and L2 to a protein, can be calculated from the free energy difference between the unbound ligands L1 and L2 in solution and the protein−ligand complex involving L1 and L2.[9,18,19]

In alchemical processes of the type represented by the horizontal arrows in Figure 1, one can distinguish two types of changes in the potential energy function U of the system. They are $\Delta U_{\mathrm{intra}}$, corresponding to energy terms restricted to the part of the system that is alchemically transformed (among others, the appropriate bond and bond angle terms belong into this group), and $\Delta U_{\mathrm{inter}}$, corresponding to the interactions between the transformed portion and the rest of the system including the solvent. To emphasize this separation, we write the total potential energy $U(\lambda)$ as a function of the coupling parameter λ that formally describes the transformation from the initial to the final state as

$$U(\lambda) = U_0 + \Delta U_{\mathrm{intra}}(\lambda) + \Delta U_{\mathrm{inter}}(\lambda) \tag{3}$$

where U_0 is the part of the energy function that is not altered by the mutation. This separation is exact for most of the empirical energy functions used in standard force fields, which

---

* Corresponding author.
[†] Department of Chemistry.
[‡] Laboratoire de Chimie Biophysique.
[§] Institut für Theoretische Chemie.

10.1021/jp981628n CCC: $18.00

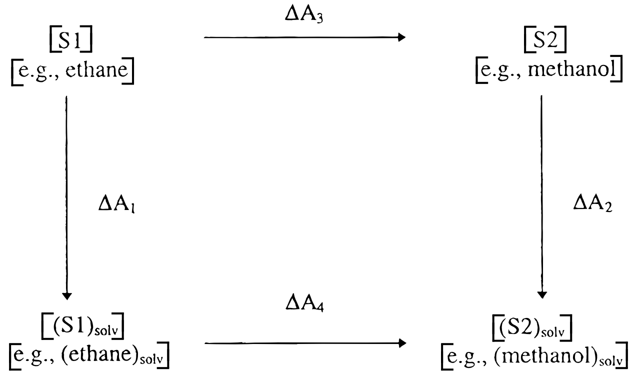

Figure 1. Thermodynamic cycle used to calculate the free energy difference of solvation between two molecules S1 and S2 (e.g., ethane and methanol).

are expressed as a sum of terms.[17,20,21] Free energy differences arising from the change in ΔU_{intra} have been referred to as *self-terms*[6] or *intragroup perturbed interactions*;[22] a detailed description of the contributions to the self-term that arise in the mutation Ile 96 to Ala in the protein barnase is given in Prevost et al.[6]

The published literature leaves unanswered a number of questions concerning the calculational methodology and the interpretation of the self-terms resulting from ΔU_{intra}. Along the two alchemical steps of a thermodynamic cycle (ΔA_3 and ΔA_4 in Figure 1), the terms in the potential function belonging to ΔU_{intra} are changed identically; i.e., the same transformation is carried out in the gas phase and in solution. This has led to the assumption that the self-terms cancel in the double free energy difference and that the contributions arising from ΔU_{intra} can be omitted,[13,23] although the simulations making this assumption were made with the full potential energy function (eq 3). Such an approach has been criticized,[24] and other workers have included ΔU_{intra} in the free energy difference formalism.[6,7,10] In general, the free energy differences resulting from the two approaches are not identical. The contribution of ΔU_{intra} to the free energy difference between S1 and S2 in solution (ΔA_4), which also contains the solute−solvent terms corresponding to ΔU_{inter} and possible coupling between the interactions, is expected to be different from that in the gas phase (ΔA_3), which contains only the intrasolute terms.

Bond and bond angle energy terms are an important part of ΔU_{intra}. Qualitatively different results concerning the importance of bonded terms in self-term contributions have been obtained in calculations of free energy differences that have included them. In a study on the free energy difference of unfolding between wild-type barnase and the I96A mutant, Prévost et al.[6] reported a large self-term, which was attributed mainly to changes in bonded interaction in the folded and unfolded state. Pearlman and Kollman have reported an "overlooked bond-stretching contribution to free energy differences" and also concluded that self-terms (which they refer to as *intraperturbed group interactions*) can be significant.[22] Nilsson and co-workers have compared the results of calculations in which contributions from bonded energy terms were included or omitted.[25,26] They found significant differences, but the results are inconclusive because of convergence problems encountered in the simulations that included the bonded energy terms in the free energy formalism. This suggests that there are methodological problems, as well as questions of interpretation, associated with the inclusion of bond length and bond angle terms in the free energy formalism. Harris and Loew recently presented some results regarding these questions and found self-terms to be of little importance;[27] similar observations were reported by Rao et al.[28] and by Sun et al.[29]

An essential question concerns the magnitude of any neglected contribution to the $\Delta\Delta A$ values when the terms arising from ΔU_{intra} are omitted. Our purpose in this paper is to focus on $\Delta\Delta A_{\text{solv}}$ and to answer the question for bond stretch and bond angle terms. In so doing we resolve some methodological problems that have arisen in the calculation of these terms. The investigation is aided by the decomposition of the total free energy difference obtained from the calculations. The so-called "component analysis" that allows one to do this has been criticized as the resulting free energy components depend on the simulation path.[30,31] We have shown elsewhere that the results of a component analysis are meaningful[19,32,33] and that the choice of path provides an additional degree of freedom that gives insights into different aspects of a free energy change. The use of component analysis to study the importance of self-terms is an extension of earlier work.[19,32,33]

We believe that an investigation of the type reported here is of considerable importance at the present time. In the past, limited computer resources have prevented fully converged calculations in some cases.[30,34] Also papers were published, as mentioned above, criticizing the significance of free energy components.[30,31,35] Since they are one of the more interesting results of free energy simulations, there seems to have been a lull in methodological developments and applications during the past few years, relative to the large number of papers published between 1987 and 1992. The results of the present paper on treating bonded terms, the development of practical approaches for solving the so-called van der Waals endpoint catastrophe,[36−38] and the availability of accurate treatments of long-range electrostatic effects[39,40] together provide a methodology for meaningful free energy simulations. When combined with sufficient computer time to obtain converged results and to do the calculation several times, free energy simulations are now ready for their rightful place in the microscopic analysis of mesoscopic systems. A good example of what can be done today is given in a study of the difference in the free energy of binding of asparagine and aspartic acid to aspartyl−tRNA synthetase.[39,41]

This paper addresses the important aspects of the role of bonded terms in free energy simulations and is organized as follows. First, a brief review of the free energy formalism and the path dependence of free energy components is given (section 2a). To analyze the role of bond stretching and bond angle bending terms in free energy simulations, a comparison is made of two ways of setting up the hybrid potential energy function (single and dual topology) (section 2b). These two simulation methodologies involve different end states. It is shown that the contributions resulting from the use of the different end states cancel from the double free energy of interest $\Delta\Delta A_{\text{solv}}$ and that they can be calculated analytically. Different (single) free energy differences and, in particular, different free energy components for the pieces of the thermodynamic cycle are obtained that provide insights into the meaning of free energy contributions from bonded terms. We then describe how to include changes in bonded terms in the free energy formalism. The appropriate method depends on whether a single or a dual topology formalism is used. It is necessary also to distinguish between changing the strength of a bond (section 2c) and deleting or adding a bond (section 2d). Finally, some observations regarding the analysis of free energies differences of solvation are presented (section 2e). A numerical example illustrating a point made in the text concerning bond angle and Urey–Bradley terms is presented in an appendix.

The theoretical considerations presented in this paper are complemented and augmented by calculations of free energy differences for selected model systems in the companion paper.[42] We report the results and interpretation of free energy differences of solvation between (i) one-dimensional diatomics, for which the partition functions can be obtained with high precision by numerical integration, (ii) between two diatomic molecules in three-dimensions, which make it possible to study the effect of altering bond parameters most clearly, and (iii) between ethane and methanol as an example of a realistic system.

2. Theory

2a. Review of the Free Energy Formalism. We begin with a brief summary of the theory used in free energy simulations. All formulas are developed in the canonical ensemble (constant volume V, temperature T, and number of particles N). Similar expressions are valid at constant pressure P, T, and N; in most cases, the two types of results should be essentially the same, although for a large volume change in an alchemical simulation, a constant pressure calculation may be more appropriate. If the potential energy function of the system is given in the form $U = U_o + \Delta U$, where ΔU is the change in potential energy relative to the reference system described by U_o, the (configurational) free energy A of the full system can be written in the form $A = A_o + \Delta A$, where ΔA is the free energy difference due to the change in the potential energy function. This is often referred to as a "perturbation", although the calculational methods used in most studies are exact, in principle. The kinetic energy contribution can be calculated analytically or ignored since it cancels in a double free energy difference.[1]

Two equivalent, exact expressions for ΔA are employed in most free energy difference simulations. The first one is commonly referred to as the thermodynamic integration method (TI)[43,44] and the second as the exponential formula (EF) (also denoted as the perturbation method or formula, even though it is, in principle, an exact expression).[45,46] In both cases, one starts by introducing a coupling parameter λ and writes the potential energy function of the system as follows

$$U(\lambda) = U_0 + \Delta U(\lambda) \tag{4}$$

We refer to this λ-dependent potential energy function as a *hybrid potential energy function* because in an alchemical simulation it makes possible the smooth transformation of the initial system to the final state by introducing an intermediate (hybrid) system. The parameter λ ranges between λ_{initial} and λ_{final}, which are commonly chosen as 0 and 1 so that $U(\lambda = 0)$ corresponds to the reference system (or $\Delta U(\lambda = 0) = 0$), and $U(\lambda = 1)$ to the perturbed system. In TI,[43,44] the free energy difference between initial ($\lambda = 0$) and final state ($\lambda = 1$) of the system is obtained as

$$\Delta A = \int_0^1 \mathrm{d}\lambda \, \langle \partial \Delta U(\lambda)/\partial\lambda \rangle_\lambda \tag{5}$$

where the integration has been performed with a variety of methods; use of the trapezoidal rule is most common. In EF, the same free energy difference is obtained as

$$\Delta A = -k_B T \ln \langle \exp[-\Delta U(\lambda = 1)/k_B \mathrm{T}\rangle_{\lambda=0} = \\ -k_B T \sum_i \ln\langle \exp[-(\Delta U(\lambda_{i+1}) - \Delta U(\lambda_\mathrm{i})/k_B T]\rangle_{\lambda_i} \tag{6}$$

In most practical applications of eq 6, the second equality is used to take into account the fact that the required ensemble averages converge only if initial and final state are not too dissimilar. In eq 6 k_B is the Boltzmann constant, T is the temperature, and the notation $\langle X \rangle_\lambda$ in eqs 5 and 6 denotes a classical statistical mechanical ensemble average in the canonical ensemble, i.e.,

$$\langle X \rangle_\lambda = [Z(\lambda)]^{-1} \int_V \mathrm{d}\mathbf{r} \, X \exp[-U(\mathbf{r},\lambda)/k_B T] \tag{7}$$

where Z is the configurational partition function and $\mathbf{r}$ denotes the $3N$ dimensional coordinate space with V designating the integration volume. The subscript λ on the ensemble average indicates that it is evaluated at a particular value of λ.

In the remainder of the paper, we use the TI formalism because it leads to a direct relation between additive terms in the potential energy function and additive terms in the free energy difference;[8,10,19,32] i.e., if ΔU is written as a sum of terms

$$\Delta U = \sum_i \Delta U_i \tag{8}$$

the free energy difference ΔA is given by a sum of terms

$$\sum_i \int_0^1 \mathrm{d}\lambda \, \langle \partial \Delta U_i(\lambda)/\partial\lambda \rangle_\lambda = \sum_i \Delta A_i \tag{9}$$

This decomposition, which has been referred to as *component analysis*, is exact and has proven to be a very useful tool for obtaining insights into the free energy change on the molecular level.[6–8,10] However, such free energy components depend on the simulation path between the initial and final state,[8,19,30–32,34,47–51] in contrast to the total free energy difference, which is a state function. To make clear this path dependence, it is useful[8,19] to consider a system with two additive terms that contribute to ΔU and to introduce separate integration variables λ and μ associated with ΔU_1 and ΔU_2, respectively; that is,

$$U(\lambda,\mu) = U_o + \Delta U_1(\lambda) + \Delta U_2(\mu) \tag{10}$$

The free energy component ΔA_1 is given by

$$\Delta A_1(\mu) = \int_0^1 \mathrm{d}\lambda \frac{\int_V \mathrm{d}\mathbf{r}[\partial \Delta U_1(\mathbf{r},\lambda)/\partial\lambda] \exp[-U(\mathbf{r},\lambda,\mu)/k_B T]}{\int_V \mathrm{d}\mathbf{r} \exp[-U(\mathbf{r},\lambda,\mu)/k_B T]} \tag{11}$$

A corresponding expression holds for $\Delta A_2(\lambda)$. Both ΔA_1 and ΔA_2 are line integrals that depend on the integration path defined by the parameters (λ, μ). The coupling between λ and μ arises from the Boltzmann factor, which involves the total potential energy $U(\lambda,\mu)$. Equations 8 and 9 correspond to what has been called the concerted linear path ($\lambda = \mu$), in which all terms in the potential function are changed in a concerted fashion corresponding to the transmutation of one system into another. Since the free energy components are not unique, care is required to establish their physical meaning in each application. The nature of the coupling between free energy components has been considered in a more general context in previous work.[8,19,32,33,51] One property of the free energy components is evident from eqs 3, 7, and 8. Only energy terms that depend on the coupling parameter give rise to free energy components. Thus, although a free energy component depends on all energy terms of the system through the Boltzmann factor, it is projected on a specific term; i.e., on $\Delta U(\lambda)$ in eq 10. This implies that different free energy components can be obtained if different hybrid potential energy functions are used to describe the mutation of interest. One of the aims of this paper is to provide insights concerning the physically meaningful interpretation of these differences.

2b. Single and Dual Topology Simulations. Two distinct ways of implementing the hybrid potential energy function (eq 4) are used in free energy calculations. We refer to them as *single* and *dual topology* methods.[52] Calculations of the free energy difference with the single or the dual topology method correspond to employing different hybrid potential energy functions; in particular, the number of degrees of freedom of the system used in the thermodynamic integration formula is different. It is necessary to understand this difference for an analysis of the effects on free energy components resulting from changes in bond and bond angle terms and the physical meaning of such contributions. In fact, different end states are present in the two methods, so even the total free energy differences are affected. However, as we show below, the same double free energy differences are obtained, which in most cases are the quantities that can be measured directly (cf. the Introduction).

Single Topology. The single topology approach, as implemented, for example, in the PERT module of CHARMM,[20] is a one-to-one mapping process where every atom in the initial state (e.g., ethane) has a counterpart in the final state (e.g., methanol). Any energy term that depends on an atom that changes in the transformation is altered so that we can write

$$U(\mathbf{r},\lambda) = U_0(\mathbf{r}) + f(\lambda)U_{\mathrm{initial}}(\mathbf{r}) + g(\lambda)U_{\mathrm{final}}(\mathbf{r}) \qquad (12)$$

where $U_0(\mathbf{r})$ is the part of the potential energy function that does not change, $U_{\mathrm{initial}}(\mathbf{r})$ contains the energy terms unique to the initial state and $U_{\mathrm{final}}(\mathbf{r})$ those unique to the final state. The variable $\mathbf{r}$ represents the system coordinates. Since the number of atoms does not change in the transformation, the coordinate space is conserved. The functions $f(\lambda)$ and $g(\lambda)$ are such that $f(\lambda=0) = 1$, $f(\lambda=1) = 0$ and $g(\lambda=0) = 0$, $g(\lambda=1) = 1$; it is common, as in the PERT module of CHARMM, to use $f(\lambda) = 1 - \lambda$, $g(\lambda) = \lambda$ (a linear dependence on λ), but other functions of λ have been employed.[37,53-55] Equation 12 is a generalization of eq 4; for a linear dependence on λ, eq 4 is obtained if one defines $U_0' = U_0 + U_{\mathrm{initial}}$, $\Delta U = U_{\mathrm{final}} - U_{\mathrm{initial}}$. Some programs (e.g., AMBER, GROMOS) implement the single topology method by mixing the parameters (force constants, charges, etc.) rather than the entire potential energy function, as in eq 12.[56-58] This difference is not important for the present analysis.

In most alchemical simulations (e.g., ethane to methanol) the actual number of atoms in the initial and final state is different. As stated above, the number of particles in the system is assumed to be fixed in single topology calculations, so that it is not the types of atoms but their number, i.e., the dimension of the configuration space, that is conserved. To accomplish this, *dummy atoms* are introduced for atoms which exist in one state and have no counterpart in the other[24] (e.g., for the transformation of ethane to methanol in an all-atom model, two of the hydrogen atoms of the methyl group, which is mutated into a hydroxyl group, are dummy atoms in the final state (see Figure 2a)). The number of atoms in a single topology simulation is equal to the larger of the number of atoms of which the initial and final system is composed. The dummy atoms have no nonbonded energy terms; i.e., no van der Waals or electrostatic terms are associated with them, but they are connected to the rest of the system through bonded terms. These bond and bond angle terms often are assumed to be the same as those in the system from which they originate as real atoms,[22,24,54] but the treatment of dihedral and improper dihedral angles seems to vary. In some applications, the bond lengths involving bonds to dummy atoms are shortened to improve the convergence of the simulations.[22,59,60] Such changes in bonded terms to dummy atoms have to be treated specifically since they can lead to nonphysical contributions to the free energy differences of interest that have to be eliminated.[60,61] More generally, the effect of differences between a system with dummy atoms and the real system it represents (e.g., methanol with two dummy atoms compared to methanol) has to be investigated and the required corrections have to be introduced.

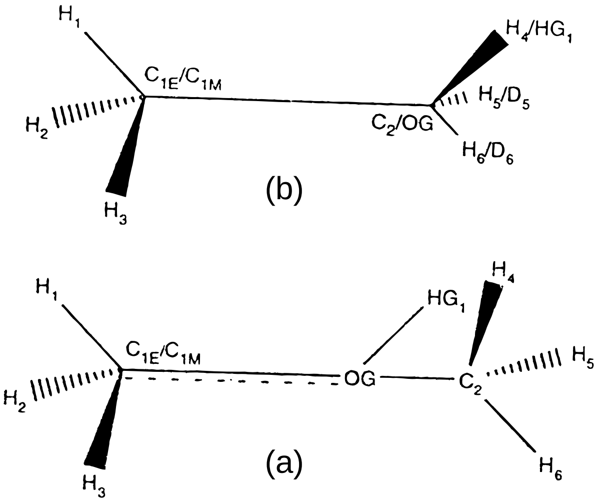

Figure 2. (a) The hybrid ethane/methanol solute used in the single topology simulations described in detail in ref 42. (b) The hybrid ethane/methanol solute used in the dual topology simulations described in detail in ref 42.

Dual Topology. The dual topology approach differs from the single topology method in that the parts of the system which are not the same in the initial and final state are defined simultaneously. Thus, the number of particles is increased, relative to the physical end states. This is exemplified for the alchemical mutation of ethane to methanol in Figure 2b. The parts of the system that change interact with the rest of the system, but not with each other. This is analogous to the use of replicas in the multiple copy simulation search method (MCSS).[62] Multiple copies have been introduced also to improve the convergence of configurational free energy simulations.[63] Examples of dual topology methods are implemented in the BLOCK and TSM modules of CHARMM. The potential energy function for the dual topology method can be written

$$U(\mathbf{r},\mathbf{r}_{\mathrm{initial}},\mathbf{r}_{\mathrm{final}},\lambda) = U_0(\mathbf{r}) + f(\lambda)U_{\mathrm{initial}}(\mathbf{r},\mathbf{r}_{\mathrm{initial}}) + g(\lambda)U_{\mathrm{final}}(\mathbf{r},\mathbf{r}_{\mathrm{final}}) \qquad (13)$$

As in the single topology method (see above), the dependence on the coupling parameter λ is often taken to be linear but it does not have to be.[37,53-55] Also different coupling parameters for different energy terms can be used. This latter capability is implemented to some degree in the TSM module of CHARMM and in a modified version of the BLOCK module.[39] The coordinates $\mathbf{r}$,$\mathbf{r}_{\mathrm{initial}}$ and $\mathbf{r}_{\mathrm{final},}$ respectively, are associated with the atoms that do not change, those that are present only in the initial state and those that are present only in the final state. Equation 13 emphasizes that the total number of degrees of freedom used in the simulation is larger than in either physical end state, i.e., the number of atoms in a dual topology simulation is the sum of the number of atoms that change between the initial and final state, plus the number of atoms that remain the same. Comparing eqs 12 and 13, we see that the essential difference between the single topology method and the dual topology method is that in the former the energy functions U_{initial} and U_{final} involve the same set of coordinates, while in the latter U_{initial} and U_{final} involve different coordinates. Usually, there are only a few more atoms in a dual topology simulation compared to the single topology calculation, because the parts of the system that do not change (e.g., the solvent and the protein) are much larger than the part that does.

For the dual topology method, the energy expression must be based on mixing potentials as in eq 13; that is, the transformation from the initial to the final state cannot be done by mixing parameters, as, for example, in the GROMOS implementation of the single topology method. This is analogous to free energy simulations based on QM or QM/MM potential functions,[64,65] where the initial and final state must be fully represented because the quantum mechanical energy function is not separable. In the dual topology methods, there is a choice concerning the scaling of the bonded interactions in the parts of the system that are modified, as well as of the bonded interactions of the modified portions with the rest of the system. It is possible to scale the bonded terms, as is done necessarily for the nonbonded terms, or not to scale them. The former approach corresponds to an ideal gas *atom* end state for the parts of the system that are changed in the dual topology method (e.g., in going from AB to A′B′, the initial state consists of the AB molecule (which interacts with solvent) and ideal gas atoms A′ and B′, while the final state consist of the ideal gas atoms A and B and the molecule A′B′ (which interacts with solvent)). This means that certain bonds and angles are broken/formed at the respective end points; in the example, the AB bond is broken and the A′B′ bond is created. A number of studies carried out with the BLOCK and TSM modules of CHARMM have used such ideal gas atom end states. If the bonded terms are not scaled, the end state corresponds to ideal gas *molecules*; in the example, the end states consist of molecules AB and A′B′, but in the initial state only AB interacts with the solvent, while in the final state only A′B′ interacts with the solvent. Whether or not the bonded terms are scaled, the parts of the system that change (AB and A′B′ in the example) do not interact with each other throughout the simulation.

In single topology methods when dummy atoms are introduced, as in the majority of practical applications, a similar ambiguity as to the choice of end state (ideal gas molecules vs ideal gas atoms) arises. If the bonded terms connecting the dummy atoms to the rest of the molecule are reduced to zero in the end states, they are transformed into ideal gas atoms; if the bonded terms to the dummy atoms are not altered, they remain part of the molecule. The latter procedure has been used in practice. However, as already mentioned, the bond lengths to dummy atoms are often changed and this can lead to effects that have to be considered.

End States. There is a conceptual question and a technical question concerned with the choice of end states. Here we focus on the conceptual aspect; the technical issue associated with making or breaking a bonded term is considered in section 2d. Since the thermodynamic cycles employed in most alchemical free energy simulations contain molecules in the initial and final states (see Figure 1), there is no need to break or make bonds. This is true for all applications of free energy simulations in which the transformation does not involve covalent bond dissociation; studies of the free energy differences of solvation, ligand binding, and protein stability are of this type. It is, therefore, possible to choose ideal gas molecules for the end state in alchemical simulations. Since the convergence is likely to be improved when bonds and bond angle terms are not made or broken, this suggests that the bonded terms be made independent of the coupling parameter in dual topology simulations. As a consequence, there would be no bond and bond angle free energy components, in contrast to simulations which scale all the energy terms associated with the parts of the system that change. The significance of this is discussed in section 2e.

The above analysis shows that single and dual topology methods generally use different hybrid potential energy functions (eqs 12 and 13) and that the end points of the simulation can be different. This is a consequence of the different philosophy behind the single and dual topology method and the requirement that the number of atoms does not change in the simulation; i.e., in the dual topology method the total number of particles corresponds to the sum of $\mathbf{r}$, $\mathbf{r}_{\mathrm{init}}$, and $\mathbf{r}_{\mathrm{final}}$, even if some of them do not interact with anything at the end points. As an illustration, the hybrid molecules used in a calculation of the free energy difference of solvation between ethane and methanol are depicted in Figure 2a for the single topology and Figure 2b for the dual topology method. Since the end points are not the same, different free energy changes for one step of a thermodynamic cycle are expected to result from the two methods. The difference in endpoints cancels from the thermodynamic cycle; therefore, identical double free energy differences, such as the free energy difference of solvation, are obtained. To show this, we address here the role of dummy atoms (which are clearly unphysical). Further differences between single and dual topology methods, as well as the choice between an ideal gas atom and an ideal gas molecule end state, are considered in section 2e.

The use of ideal gas molecule end states raises the problem that the endpoints in the simulation do not necessarily correspond to the real system. In single topology simulations when dummy atoms are required, there are additional atoms (the dummy atoms) bonded to the system of interest. Similarly, in dual topology simulations the parts of the molecule which at the respective end points do not interact with the rest of the system are present and bonded to the system. Such fragments, which usually have only harmonic bond and bond angle terms, do not affect the free energy differences of interest; dihedral angles to dummy atoms (which are not harmonic terms) can always be turned off, as is the case in the calculations reported in the companion paper. To show that the dummy atom terms cancel, we consider a system that differs from the *physical* system it attempts to mimic by one dummy atom connected via a harmonic bond stretching term and a harmonic bond angle term; at the end state the dummy atom has no nonbonded interactions. By methods analogous to those used in deriving Jacobian factors,[61,66] the contribution from the dummy atom can be factored out in the configurational partition function, which can be written as

$$Z = Z' \int \mathrm{d}r\, r^2 \exp(-\beta K(r - r_o)^2) \int \mathrm{d}\theta \sin\theta \times \exp(-\beta K_\theta(\theta - \theta_o)^2) \quad (14)$$

Here Z' is the partition function of the real system; the two integrals correspond to the harmonic bond and bond angle term of the dummy atom. They add a constant term in the parallel pieces of the thermodynamic cycle of interest and, consequently, cancel from the double free energy difference. Furthermore, if one wanted to know a single free energy difference and needed to correct for the presence of the dummy atoms, the respective terms could be calculated on the basis of eq 14. If anharmonic terms are used in the force field, the first part of the simulation could change the anharmonic terms involving atoms that become dummy atoms into fitted harmonic terms. Whether the resulting harmonic energy terms involving the dummy atoms are removed depends on the choice of atom vs molecule end state. Equation 14 can be used to account for the contributions of the harmonic terms.

Dual Topology in Single Topology Framework. An interesting application of the theory involves attempts to mimic a dual topology calculation with a program that includes only the single topology method. This has been proposed and applied by Pearlman[52] and more recently discussed by Sun et al.[29] Similar to the dual topology method, the parts of the system which are not the same in the initial and final state were defined simultaneously. In the initial state, all atoms representing the final state were dummy atoms and vice versa. The configuration space was, therefore, the same as in the dual topology method, i.e., the total number of degrees of freedom was equal to $\mathbf{r} + \mathbf{r}_{\text{init}} + \mathbf{r}_{\text{final}}$. However, the hybrid potential function was not the same as given by eq 13. The nonbonded parameters, i.e., the van der Waals parameters ϵ and σ, as well as the charges q, of the physical atoms were scaled as a function of the coupling parameter. This can be written as

$$U(\mathbf{r},\mathbf{r}_{\text{init}},\mathbf{r}_{\text{final}}) = U_{\text{o}}(\mathbf{r}) + U_{\text{init}}(\mathbf{r},\mathbf{r}_{\text{init}}, (1-\lambda)\sigma_{\text{init}}, (1-\lambda)\epsilon_{\text{init}}, (1-\lambda)q_{\text{init}}) + U_{\text{final}}(\mathbf{r},\mathbf{r}_{\text{final}}, \lambda\sigma_{\text{final}}, \lambda\epsilon_{\text{final}}, \lambda q_{\text{final}}) \quad (15)$$

One sees from eq 15 that similar to a standard dual topology method (eq 13) the potential energy terms describing the initial and final state act on different coordinates. However, the interaction parameters, rather than the energy terms, were scaled by λ. Bond and bond angles involving dummy atoms, i.e., the respective other half of the system were not changed; this corresponds to the ideal gas molecule end state discussed above. So far this approach has been applied only to the model calculation of changing ethane into ethane; see ref 52, although not all technical details are made clear in that paper, (e.g., the treatment of the dihedral angle terms is not specified explicitly). Sun et al.[29] utilized the idea to make clear the difference between single and dual topology simulations, although it should be kept in mind that these "dual topology" calculations are not identical to the use of the BLOCK or TSM module of CHARMM. Nevertheless, eq 15 presents an interesting ansatz for combining elements of parameter mixing[56−58] and the dual topology method.

2c. Contribution of Bonding Terms in Alchemical Free Energy Simulations. We now turn to the question of calculating the contributions from changes in bond stretching and bond angle bending terms in free energy simulations. Analysis of the bonded energy terms provides a way of determining the physical meaning of the free energy components associated with them. We show that the bond or bond angle free energy components have vibrational, potential-of-mean-force type and Jacobian factor contributions. In what follows, we present the formulation for bond length terms and then briefly describe the results for bond angles.

Vibrational Contributions. We consider first the free energy differences that arise from changing a one-dimensional harmonic oscillator into another in the absence of other interactions (i.e., in the gas phase). In classical statistical mechanics, it is straightforward to calculate analytically the partition function for a harmonic oscillator and from it the free energy difference that is introduced by changing the force constant and the bond length.[46] It is useful for the present analysis to do the corresponding calculation in the TI formalism. We employ a single topology hybrid potential energy function of the type used in AMBER or GROMOS. It has the form

$$\Delta U(x,\lambda) = [(1-\lambda)K_{\text{i}} + \lambda K_{\text{f}}]\,\{x - [(1-\lambda)x_{0,\text{i}} + \lambda x_{0,\text{f}}]\}^2 \quad (16)$$

Here x is the position of the particle, K is the force constant, x_{o} denotes the equilibrium bond length, and i and f stand for initial and final state, respectively. To calculate the free energy difference, we need the ensemble average $\langle \partial U/\partial\lambda\rangle_\lambda$, as in eq 5. Introducing the shorthand $\Delta x = x - [(1-\lambda)x_{0,\text{i}} + \lambda x_{0,\text{f}}]$, we obtain

$$\frac{\partial \Delta U}{\partial \lambda} = (K_{\text{f}} - K_{\text{i}})\Delta x^2 + [(1-\lambda)K_{\text{i}} + \lambda K_{\text{f}}]2\Delta x(x_{0,\text{f}} - x_{0,\text{i}}) \quad (17)$$

The ensemble average $\langle \partial U/\partial\lambda\rangle_\lambda$ as defined in eq 7 can be calculated analytically. It is equal to

$$\left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda = \frac{k_{\text{B}}T}{2}\,\frac{K_{\text{f}} - K_{\text{i}}}{K_i + \lambda(K_{\text{f}} - K_{\text{i}})} \quad (18)$$

Substitution into eq 5 yields

$$\Delta A_{\text{vibr}} = \frac{k_{\text{B}}T}{2}\ln(K_{\text{f}}/K_{\text{i}}) \quad (19)$$

which is the expected result for the free energy change when a one-dimensional harmonic oscillator is transformed into another. It is assumed that the particle masses do not change so that there is no kinetic energy contribution.[46] As can be seen from eq 19, the free energy difference is independent of the equilibrium bond length x_{o}; i.e., the only contribution arises from the change in the force constant. Qualitatively, this result corresponds to the fact that the free energy change arises from the entropy, which depends on the "size" of the configuration space available to the oscillator and not on where it is located (as determined by x_{o}). Since a change in force constant alters the vibrational frequency of the oscillator, we refer to this type of free energy contribution as *vibrational*. To do the corresponding calculation with a molecular dynamics simulation, $\langle \partial\Delta U/\partial\lambda\rangle_\lambda$ is evaluated from a trajectory at each λ; i.e., from eq 17, it is necessary to determine

$$\left\langle \frac{\partial \Delta U}{\text{d}\lambda} \right\rangle_\lambda = (K_{\text{f}} - K_{\text{i}})\,\langle \Delta x^2\rangle_\lambda + [(1-\lambda)K_{\text{i}} + \lambda K_{\text{f}}]\,2\langle \Delta x\rangle_\lambda\,(x_{0,\text{f}} - x_{0,\text{i}}) \quad (20)$$

eq 20, which was first given by Simonson and Brünger,[49] makes clear how to include changes in bond terms in the TI formalism. It is generally straightforward to obtain accurate values for the required ensemble averages $\langle \Delta x^2\rangle_\lambda$ and $\langle \Delta x\rangle_\lambda$ (see the companion paper and ref 61). The excellent convergence behavior in such a TI calculation contrasts with problems that can arise when EF is used. Pearlman and Kollman[22] reported severe convergence problems when attempting to calculate the free energy change between two harmonic oscillators with EF. The difficulty arises from an interesting difference between TI and EF. In TI (cf. eq 5) the time averages (as in eq 20) are evaluated for the value of λ at which the simulation is performed.

By contrast, in EF (as in eq 6) averages of the form $\langle \exp(-\beta \Delta U(\lambda_{i+1}) \rangle_{\lambda i}$ have to be determined; i.e., the exponential of the energy function at λ_{i+1} must be evaluated for configurations sampled at λ_i. This can lead to very slow convergence for bonds if the regions sampled for $\Delta U(\lambda_{i+1})$ and $\Delta U(\lambda_i)$ do not overlap well. Severance et al. have recently shown that this limitation can be overcome in EF by correcting for the change in equilibrium bond length r_o as a function of the coupling parameter;[67] i.e., they observed that the potential energy function of a bond term can be rewritten as

$$U(\lambda) = K(\lambda)[x(t) - x_o(\lambda)]^2 = K(\lambda)[\Delta x(\mathrm{t})]^2 \qquad (21)$$

if $x(t)$ (the instantaneous bond length) is written as

$$x(t) = x_o(\lambda) + \Delta x(t) \qquad (22)$$

(cf. eq 7 of ref 67). The convergence problem encountered by Pearlman and Kollman[22] occurs exactly because of the λ-dependence of x_o: In the EF formalism, one needs to compute the difference (see eq 6)

$$U(\lambda_{i+1}) - U(\lambda_i) = K(\lambda_{i+1})(x_{\lambda_i}(t) - \mathrm{x}_o(\lambda_{i+1}))^2 - K(\lambda_i)(x_{\lambda_i}(t) - x_o(\lambda_i))^2 \qquad (23)$$

The notation $x_{\lambda_i}(t)$ in eq 23 makes clear that the instantaneous value of the bond length is calculated with a potential corresponding to λ_i, in general, $x_{\lambda_i}(t)$ and $x_o(\lambda_{i+1})$ will be quite different; hence, the first term in eq 23 will be large so that the convergence will be slow upon insertion into eq 6. To avoid this problem, Severance et al. suggest taking into account the λ-dependent part of $x(t)$ by replacing $x_{\lambda_i}(t)$ by $x_{\lambda_{i+1}}(t) = x_o(\lambda_{i+1}) + \Delta x(t)$, based on the second expression in eq 21. In this case, eq 23 becomes

$$U(\lambda_{i+1}) - U(\lambda_i) = [K(\lambda_{i+1}) - K(\lambda_i)]\Delta x(t)^2 \qquad (24)$$

which is well behaved. Although it is not clear that eq 24 is exact, its use leads to correct results, at least within certain error limits, as demonstrated by the test calculations presented in ref 67.

External Forces or Potential-of-Mean-Force Type Contributions. Although in the isolated harmonic oscillator the free energy depends only on the force constant and not on the bond length, this is not true when external forces are present; their effect is analogous to that which arises in transforming an anharmonic oscillator. To examine the influence of external forces (e.g., the harmonic oscillator interacts with a field or is immersed in solvent), we consider the case in which the force constant is unchanged but the equilibrium bond length is increased or decreased. Equation 20 simplifies to $\langle \partial \Delta U/\partial \lambda \rangle_\lambda = 2K\langle \Delta x \rangle_\lambda\ (x_{0,\mathrm{f}} - x_{0,\mathrm{i}})$. For an isolated harmonic oscillator (e.g., in the gas phase), $\langle \Delta x \rangle_\lambda$ is zero for all values of λ since the potential is symmetric about $x_o(\lambda)$. This is in accord with the theoretical result that the change in bond length does not affect the free energy (eq 19). In the presence of interactions (e.g., in solution or in the presence of an external electric field for a dipolar oscillator), $\langle \Delta x \rangle_\lambda$ can deviate from zero and a nonzero contribution to the free energy difference can result. The free energy change is introduced by the altered average bond length induced by the presence of external forces; i.e., it reflects the change in potential of mean force acting on the oscillator as a function of its bond length x_o. To show this we transform to the equilibrium bond length x_o, instead of λ, as the coupling parameter. With $x_o(\lambda{=}0) = x_{\mathrm{o,i}}$ and $x_o(\lambda{=}1) = x_{\mathrm{o,f}}$ the free energy difference (eq 5) becomes

$$\Delta A = \int_{x_{\mathrm{o,i}}}^{x_{\mathrm{o,f}}} \mathrm{d}x_\mathrm{o}\ \langle \partial \Delta U/\partial x \rangle_{x_\mathrm{o}} \qquad (25)$$

where $\langle \partial \Delta U/\mathrm{d}x \rangle_{\mathrm{x_o}}$ is the average force acting along the bond. Even if the parameters describing the interactions between the harmonic oscillator (“solute”) and external forces (“solvent”) are constant, i.e., in a physical system this would mean that van der Waals parameters and charges remain the same, the average solute−solvent interaction energy and, consequently, $\langle \partial \Delta U/\partial x \rangle_{\mathrm{x_o}}$ can change as a function of x_o. Equation 25 is the definition of the change in potential of mean force for the interaction of two particles as a function of their distance x_o.[46] We, therefore, refer to the free energy contributions to bonded terms arising from external forces as *pmf-type* contributions.

Frequently, bonds are kept fixed by means of constraints, of which SHAKE[68] is the one that is most commonly used in molecular dynamics simulations. In the presence of such constraints, there is no *vibrational* contribution to the free energy difference since the bond degree of freedom has been removed. This is an approximation, though the neglected effect is likely to be small. However, *pmf-type* contributions still exist and have to be taken into account. The general theory of introducing holonomic constraints in free energy difference calculations is summarized in section 4.4 of ref 48. Here we only consider the case of constraining bond lengths; in this case the pmf-contribution can be obtained as follows (cf. section 4.5 of ref 48). If the equilibrium bond length changes as a function of the atom types in an alchemical transformation, the target distance of the constraint becomes a function of the coupling parameter λ. We use a constraint potential of the form, cf. ref 48,

$$U_{\mathrm{cons}}(\lambda) = \mu(\lambda)\delta(x_o(\lambda)) \qquad (26)$$

where $\delta(x_o(\lambda))$ is the constraint that depends on the target distance $x_o(\lambda)$

$$\delta(x_o(\lambda)) = x^2 - (x_o(\lambda))^2 = 0 \qquad (27)$$

with $x_o(\lambda) = (1 - \lambda)x_{\mathrm{o,i}} + \lambda x_{\mathrm{o,f}}$, and $\mu(\lambda)$ is the Lagrangian multiplier that is determined from the constraint equation in the standard manner (e.g., by the use of SHAKE).[68] Using eqs 5, 26, and 27, the resulting free energy change for a rigid diatomic molecule in one dimension is

$$\Delta A = \int_0^1 \mathrm{d}\lambda\ \langle \partial U_{\mathrm{cons}}(\lambda)/\partial \lambda \rangle_\lambda = -2\int_0^1 x_o(\lambda)[x_{\mathrm{of}} - x_{\mathrm{oi}}]\langle \mu \rangle_\lambda\ \mathrm{d}\lambda \qquad (28)$$

where the term involving the derivative $\partial \mu/\partial \lambda$ cancels because of eq 27.

A number of publications have compared methods for calculating such “constraint” or “pmf” corrections; a good overview can be found in ref 60. In addition, this contribution can also be calculated based on the generalization of eq 28 to more than one constraint as first pointed out by van Gunsteren and co-workers;[48] apparently, this possibility is not well-known. The derivative of the free energy due to N_c bond length constraints depending on the coupling parameter λ is

$$\partial A_c/\partial \lambda = \langle \partial U_{\mathrm{cons}}(\lambda)/\partial \lambda \rangle_\lambda = -2\sum_{k=1}^{N_c} \langle \mu_k \rangle_\lambda\ x_{o,k}(\lambda)\ [x_{\mathrm{of,k}} - \mathrm{x}_{\mathrm{oi,k}}] \qquad (29)$$

Clearly, only those bonds whose equilibrium bond length is changed need to be considered. Equation 29, which is identical to eq 4.5.9 of ref 48, is only valid for bond length constraints. It forms the basis for the method to compute constraint corrections used in the PERT module of CHARMM C24b1 and subsequent versions of the program (S. Boresch, unpublished).

Jacobian Factor Contributions. There is one more contribution that can appear as part of the bonded free energy components in single topology simulations for flexible bonds or as part of the constraint correction for constrained bond terms. Since it has been analyzed in detail recently and shown to involve Jacobian factors,[61] we describe it here only briefly. The same contributions have been formulated somewhat differently in earlier work and referred to as *moment of inertia corrections*[59] or *dynamic stretch free energies*.[60] The simplest case where such a contribution arises is in a three-dimensional harmonic oscillator.[69] We associate the three translational degrees of freedom with the Cartesian coordinates of the first atom and describe the position of the second atom in polar coordinates relative to the first atom. The configurational partition function of the system expressed in terms of these coordinates is

$$Z(\lambda) = V\,4\pi \int_0^\infty \mathrm{d}r\, r^2 \exp[-\beta K(\lambda)(r - r_o(\lambda))_2] \quad (30)$$

Here, V is the volume of the system and results from the integration over the coordinates of the first particle, and the factor 4π arises from the integration over the two polar coordinates ϑ and φ describing the orientation of the second particle with respect to the first. The remaining integration in the configurational partition function is over the distance r between the two particles. This is analogous to the integration for a one-dimensional oscillator, except for the presence of the factor r^2, which comes from the Jacobian factor $r^2 \sin \varphi$ that arises in the transformation of the Cartesian coordinates of the second particle to polar coordinates with respect to the first. As indicated, the limits of integration are between zero and ∞, rather than between $-\infty$ and $+\infty$ for the one-dimensional case. For most values of the force constant K, the distance r between the two particles is limited to a very narrow range near $r_{o,}$ so that the rigid rotor (RR) approximation can be used; it corresponds to the limit $K \rightarrow \infty$ in eq 30. In this limit, r^2 is replaced by $r_o{}^2$ which can be taken out of the integral. The free energy difference between two isolated three-dimensional harmonic oscillator is then

$$\Delta A = -k_B T \ln(r_{of}/r_{oi})^2 + (k_B T/2) \ln (K_f/K_i) = \Delta A_J + \Delta A_{vibr} \quad (31)$$

The first term in eq 31 is the free energy difference due to the Jacobian factor (ΔA_J); the second term is the vibrational free energy contribution given in eq 19. For the general case, ΔA_J can be calculated analytically within the RR approximation for any change in molecular geometry by use of the expression

$$\Delta A_J = -k_B T \ln(\prod_i J_{iS}{}^f / J_{iS}{}^i) \quad (32)$$

where the J_{iS} are the Jacobian factors for the transformation from Cartesian to internal coordinates S.[61,66] Jacobian free energy contributions arise both for flexible and constrained bond terms. (In the appendix of ref 61 it is shown that the formalism outlined here is to a very good approximation also valid for constrained bond terms.) If the free energy contribution from a bond term is calculated with eq 20, any contribution from Jacobian factors is included; the same is true for constraint corrections calculated with eq 29. The PMF correction of Pearlman and Kollman[22] and the PF method suggested by Pearlman[70] compute the pmf-type contribution to the free energy for systems with constrained bond terms without including the Jacobian factor contribution.[60,61] Consequently, these methods avoid the need to correct for unphysical Jacobian factors, which arise if bond lengths involving dummy atoms are shortened. The contribution due to such a change in geometry is unphysical, since dummy atoms are an artifact of the simulation methodology,[60,61] see the subsection on dummy atoms below. On the other hand, methods that do not include the Jacobian factor contribution in the free energy difference may omit physical contributions that can arise; an example is the change of a carbon−carbon bond into a carbon−oxygen bond. Most such physical contributions tend to be small and any omission can be corrected by the analytical techniques described in ref 61.

For a change in the equilibrium bond length and force constant of a bond term (in three dimensions) for a molecule interacting with a solvent environment, the contributions to the free energy are

$$\Delta A = \Delta A_{PMF} + \Delta A_J + \Delta A_{vibr} \quad (33)$$

The terms ΔA_J and ΔA_{vibr} are defined in eq 31; ΔA_{PMF} is given by eq 25 or eq 28. An important difference from the one-dimensional case discussed earlier is the dependence of ΔA on the change in bond length due to the Jacobian factor contribution, even in the absence of a pmf-type term. It can be evaluated separately as long as the RR approximation is valid. This is not a necessary condition for correct simulations, which always give the full ΔA term, as in eq 33. The analytic techniques discussed in this section that depend on the RR approximation are required only to dissect the total bond free energy into contributions, not to calculate it in a simulation. We showed recently[61] that use of the RRHO approximation may also allow one to omit one leg of a thermodynamic cycle for some systems.

Coupling Between Contributions. Within the RRHO approximation the free energy difference $\Delta A = \Delta A_{bond}$ arising from a change in bond parameters (force constant and/or equilibrium bond length) in single topology simulations can be separated into three contributions, as reflected in eq 33. It was shown that the potential of mean force type contribution is caused by the change in nonbonded solute−solvent interactions, although it is projected on to the bond free energy component. In dual topology simulations, this contribution appears as a nonbonded free energy component. If the RRHO approximation is significantly in error for a given system, ΔA_{bond} obtained in a simulation will contain contributions from coupling between the three terms of eq 33.

The energy terms associated with flexible bonds (ΔA_{vibr} and ΔA_J) may be influenced by the interaction with the environment (e.g., solvent). To illustrate this, we consider the case of diatomic, polar molecules solvated in water. The nonbonded interactions may affect the vibrational properties of the system; thus, if the free energy difference between two such species were computed, a different ΔA_{vibr} than in the gas phase would be obtained. Similarly, the interaction with the environment may change the average bond length of the solute; in a free energy difference calculation this would lead to a ΔA_J that is slightly different from the result expected from eq 32. Both types of coupling have been observed, e.g., refs 71 and 72. As the coupling present in the solvated state does not occur in the gas phase, there is a self-term contribution to the double free energy difference of solvation, which is of interest in the context of this study. However, such effects are expected to be small in most cases. For example, in a detailed quantum mechanical study of the effect of hydration on *N*-methylacetamide, a maximum increase in the N−H bond length of 0.01 Å was found due to hydrogen bonds;[71] this would change the Jacobian free energy contribution by 0.006 kcal/mol. Further, although the free energy methodologies discussed here include coupling contributions, the standard force fields commonly used for solution studies[17,20,21] may not be sufficient to describe such effects correctly. For example, De Souza et al. investigated frequency shifts induced by a simple solvent (liquid argon) for a diatomic molecule (Br_2) by computer simulation.[73] To correctly reproduce quantum mechanical results, they found it necessary to use an anharmonic (cubic) oscillator term, as well as a solute−solvent potential that depends on the bond length. Neither of these are included in the force fields most often used in macromolecular simulations.[17,20,21]

Coupling between nonbonded interactions and the vibrational and Jacobian factor contributions can be determined from explicit comparisons of simulation results using flexible bond terms with those where bonds are constrained. The main difference between constrained and flexible bond terms is the absence of vibrational free energy contributions in the former; therefore, any such coupling is missing from simulations using constrained bond terms. If bond lengths are fixed by constraints, the RR approximation is fulfilled for these degrees of freedom, and eq 32 describing the Jacobian factor contribution is exact in this case. These points are illustrated by the comparison between rigid and flexible systems made in the companion paper.

Considerations Regarding Dummy Atoms. Free energy contributions from changes in bond terms that involve dummy atoms merit a special discussion as dummy atoms are a necessary element of the single topology method.[24] If the bond term is left unchanged from its physical counterpart; i.e., the same bond length and force constant is used for dummy atoms, no free energy contribution arises. However, as mentioned earlier, bond lengths to dummy atoms are often shortened to reduce van der Waals type endpoint problems. In this case, a pmf-type and a Jacobian factor contribution result. As discussed above, the latter is unphysical since the change in bond length does not correspond to a physical process,[60,61] and the corresponding free energy contribution should not be included in the final result. The Jacobian factor contribution cancels from a thermodynamic cycle if the same change in bond length is introduced in both parts of the cycle. Alternatively, it can be calculated analytically to a good approximation with the Jacobian factor formalism (eq 32) and subtracted as appropriate.

The pmf-type contribution from a change in bond length to dummy atoms is a physical contribution and must not be omitted. When an atom becomes a dummy particle, all its nonbonded interactions with the rest of the system are removed. This can be done by turning off the nonbonded terms without changing the bond length. In this case, only nonbonded free energy components result. Alternatively, to achieve quicker convergence the bonds are shrunk as the nonbonded terms are decreased.[54,60] In this case, different nonbonded free energy components plus a pmf-type bonded contribution will result. The pmf-type bond contribution is part of the free energy difference of interest as it accounts for part of the loss of nonbonded interactions with the rest of the system. Provided that unphysical Jacobian factor contributions have been removed, the sum of these two contributions (nonbonded plus pmf) is equal to the nonbonded contribution from the calculation with constant bond lengths.

Bond Angle Contributions. Bond angles are also described by harmonic potentials in the typical molecular mechanics force fields.[17,20,21] The Urey−Bradley terms used in the new CHARMM22 all-atom parameters are also harmonic[73] and act on 1−3 distances (two atoms bonded to the same atom). Thus, all results and conclusions obtained for bond-stretching terms apply essentially unchanged to bond angle and Urey−Bradley terms.

To illustrate this point, we consider the partition function of a nonlinear, triatomic molecule. It has the form[66]

$$Z(\lambda) = V8\pi^2 \int_0^\infty \mathrm{d}r\, r^2 \exp(-\beta K(\lambda)(r - r_\mathrm{o}(\lambda))^2) \int_0^\infty \times \mathrm{d}r'\, r'^2 \exp(-\beta K'(\lambda)(r - r_\mathrm{o}'(\lambda))^2) \int_0^\pi \mathrm{d}\theta \sin\theta \times \exp(-\beta K_\theta(\lambda)(\theta - \theta_\mathrm{o}(\lambda))^2) \quad (34)$$

where, for the triatomic molecule ABC, r is the distance between B and A, r' is the distance between C and B, and θ is the angle $\angle$CBA. The factor $8\pi^2 V$ results from the integration over the degrees of freedom corresponding to the overall translation and rotation of the molecule. The first two integrals are the partition functions for the two bond stretching terms; they are identical in form to the integral in eq 30. The third configurational integral is for the angle term. The only difference in form relative to the bond stretching terms is in the Jacobian, which equals $\sin\theta$ instead of r^2. If the system were described by two bond-stretch terms in r and r' plus a Urey−Bradley term between A and C in r'', then eq 34 would change slightly. In this case, the limits of the integral go from 0 to ∞ for r and r' as before and from 0 to $r + r'$ for r'', and the Jacobian factor for the (full) system is $8\pi^2 Vrr'r''$.[66] The Jacobians in all three integrals in eq 34 change, reflecting that the 1−3 distance is not independent of r and r'.

In TI, the free energy differences from a change in bond angle θ and/or force constant K_θ can be calculated according to eq 20 with Δx replaced by $\Delta\theta$, where $\Delta\theta = \theta - \theta_\mathrm{o}(\lambda)$. Within the RRHO approximation, the resulting free energy difference contains a vibrational contribution for a change in force constant (eq 19), a Jacobian factor contribution for a change in equilibrium bond angle, and, in the presence of solvent, a pmf-type contribution if the equilibrium bond angle is changed. The Jacobian factor contribution is proportional to $\ln(\sin\theta_\mathrm{f}/\sin\theta_\mathrm{i})$[61] (unless one chooses the description using the Urey−Bradley 1−3 distance), which is usually small, as most equilibrium bond angles are relatively similar; for the same reason, pmf-type contributions are expected to be small. In contrast to bond lengths, bond angle terms are usually not constrained by SHAKE,[75] so the calculation of a constraint correction is not considered here. Coupling between vibrational and pmf-type contributions may be more important than for bond stretching terms as the force constants of bond angle terms are weaker.

A difficulty arises if a bond angle degree of freedom is determined simultaneously by an angle term in θ and a Urey−Bradley term in the 1−3 distance, which is frequently the case in the new CHARMM all-atom force field.[74] The full details, including a numerical example, are given in the appendix. Considering again the nonlinear triatomic molecule just used as an example, one sees that in this case the *three* degrees of freedom are described by *four* parameters; i.e., r, r', r'', and θ are not independent from each other. To obtain analytical formulas, this has to be taken into account properly. One possibility is to keep the two bond lengths and the bond angle as the independent (nonredundant) internal coordinates. Thus the Jacobian factor for the third atom remains $r'^2 \sin\theta$ as before, but the vibrational contribution from the bond angle force constant has to be replaced by an effective term that accounts for the influence of the Urey−Bradley term. This is equivalent to approximating the combined bond angle and Urey−Bradley potential energy by

$$U_{\text{angle}} + U_{\text{UB}} \approx K_{\theta,\text{eff}}\,(\theta - \theta_o)^2 \tag{35}$$

As shown in the appendix, $K_{\theta,\text{eff}}$ is a function of K_θ, K_{UB}, r_o, r_o', and θ_o. We stress that these considerations are not important for free energy difference simulations based on molecular dynamics, which correctly include all contributions, but are relevant only to the decomposition of bonded free energy contributions for purposes of analysis and interpretation based on eq 33.

2d. Breaking or Making a Bond. We now consider the possibility of breaking or making a bond; i.e., the situation where K_i or K_f goes to zero. Such a process is required if one turns off all energy terms (including bond terms) to dummy atoms (single topology), or if one scales nonbonded and bonded interactions alike in dual topology simulations, (i.e., if one chooses an ideal gas atom end state, cf. section 2b). It is clear that the expression for the free energy difference (eq 19) fails in this case if it is used directly. Similarly, when K_f is set equal to zero in eq 18, one obtains

$$\left\langle \frac{\partial U}{\partial \lambda} \right\rangle_\lambda = -\frac{k_B T}{2}\frac{1}{1-\lambda} \tag{36}$$

which, upon integration, leads to an expression that diverges logarithmically as λ goes to 1. In both eqs 18 and 19 the fact that the system is confined to a volume V which would lead to a finite value when a bond is broken is not taken into account. To determine what happens, we consider eq 30, the partition function of a three-dimensional harmonic oscillator, for $K(\lambda) \rightarrow 0$. Clearly, the RR approximation is *not* applicable in this case; i.e., one cannot set $r^2 = r_o^2$ and take the term out of the integral. Instead, since $\exp[-\beta K(\lambda)(\mathrm{r} - \mathrm{r}_o(\lambda)]^2$ goes to unity as $K(\lambda) \rightarrow 0$, eq 30 becomes

$$Z(\lambda \rightarrow 0) = V\,4\pi \int_0^r \mathrm{d}r\, r^2 = V^2 \tag{37}$$

where the upper limit of integration determines the system volume. This is the expected result for the configurational partition function for two noninteracting particles confined to a volume V. For the calculation of the equilibrium constant between a diatomic and the two separated atoms in a volume V, the bond energy of the diatomic would have to be taken into account in the standard manner.[46]

Use of the formulation just described for bond breaking in simulations leads to difficulties because the particles would have to sample the full volume of the system. In most cases, this leads to configurations which are unrealistic for any value of the coupling parameter other than zero or one, and adequate sampling is not realizable in simulations of finite length. Erratic results were obtained in single and dual topology simulations that attempted to break bonds in this way (S. Boresch & M. Karplus, unpublished results). It seems likely that this is also the reason for the convergence problems observed in refs 25 and 26. Since the TSM module of CHARMM used in these calculations is a dual topology method the inclusion of the coupling parameter for bond and bond angle terms led to breaking and forming of bonds and bond angles.

One solution is to consistently use an ideal gas molecule end state, i.e., not to scale any bond and bond angle energy terms. The appropriateness of such an approach for the computation of double free energy differences of physical interest defined by thermodynamic cycles was already shown in section 2b, subsection End States. Results identical to those from a single topology simulation or a dual topology simulation using an ideal gas atom end state are obtained. However, this is not true for the individual single free energy differences since the end states involved in the simulations are not the same. In addition, the free energy components will be quite different. There are no bonded contributions (as in a single topology simulation, section 2c) because these energy terms do not depend on the coupling parameter if an ideal gas molecule end state is employed. The respective contributions to the double free energy difference, however, are not omitted, but enter indirectly through the Boltzmann density as part of nonbonded free energy components. This is discussed further in section 2e.

An alternative approach to the problem of creating or annihilating a bond when an ideal gas atom end state is used is to introduce a criterion for determining when a bond is broken. The exact expression for the configurational free energy of a bond described by a harmonic potential $U = K/2(x - x_o)^2$ is[46]

$$A = -\frac{k_B T}{2}\ln\frac{2\pi k_B T}{K} \tag{38}$$

Here, k_B is the Boltzmann constant, T the temperature, and K the force constant. To obtain this result we set K_i in eq 18 equal to zero (this corresponds to the formation of a bond with force constant $\mathrm{K} = K_f$) and for the integration introduce a lower limit $\lambda = \epsilon$ (rather than $\lambda = 0$), where the limiting value corresponds to the use of a criterion for a broken bond. This leads to the integral

$$A = \frac{k_B T}{2}\int_\epsilon^1 \mathrm{d}\lambda/\lambda \tag{39}$$

The value of ϵ is chosen such that eq 39 yields the same result as eq 38, i.e.,

$$\epsilon = 2\pi k_B T/K_f \tag{40}$$

The case of breaking a bond can be treated analogously by carrying out the integration from zero to $1 - \epsilon$, where $\epsilon = 2\pi k_B T/K_i$. These limits for λ can form the basis for calculating the free energy difference of bond creation and annihilation in a molecular dynamics simulation by use of eq 20, for example.

The bound ϵ or $(1 - \epsilon)$ that yields the correct free energy has a simple physical meaning. In dual topology free energy simulations the effective force constant in the simulation is given by $K(\lambda) = \lambda K$ or $(1 - \lambda)K$ when a linear dependence of the coupling parameter is used. Thus, $\lambda = \epsilon$ or $1 - \epsilon$ corresponds to the limiting (lowest) value of the force constant for which the bond is considered to still exist. This value, denoted as K_ϵ, is found for both bond formation and annihilation to be equal to $K_\epsilon = 2\pi k_B T$. Inserting K_ϵ into eq 38 yields a zero free energy since the expression in the logarithm, which is the configurational partition function, equals unity. The configurational partition function has the dimensionality of $[L]$ in one dimension (or $[V]$ in three dimensions) so that the bound on λ corresponds to a limiting force constant K_ϵ at which the particle samples a unit length (or volume). The omitted range of integration, $[0,\epsilon]$ for bond forming and $[1-\epsilon,1]$ for bond breaking, accounts for the fact that the free particle has access to the full length L (or volume V) of the system, not just a unit length (or volume). Since this is a constant contribution, the choice of ϵ and $1 - \epsilon$ as the thresholds at which a bond is formed or broken is appropriate.

For a harmonic oscillator in three dimensions, eq 38 is replaced by

$$A = -k_BT \ln (4\pi^2 r_o^{\,2}) - (k_BT/2)\ln(2\pi k_BT/K) \qquad (41)$$

where r_o is the equilibrium bond length and the first term arises from the Jacobian factor. We choose to exclude the Jacobian factor contribution (first term) from consideration as it can always be dealt with analytically to a very good approximation. The second term of eq 41 is identical to eq 38. However, the interpretation of $K \rightarrow K_\epsilon$ changes slightly. The volume factor for the second atom of a three-dimensional harmonic oscillator is $4\pi r^2$ (the Jacobian) times a thickness which is a measure of the average vibrational amplitude of the bond; i.e., $(2\pi k_BT/K)^{1/2}$; for $K \rightarrow K_\epsilon$, this thickness is equal to unity. A discussion of the relationship between the configurational partition function and the volume accessible to each atom of the molecule is given in refs 61 and 66.

The introduction of a cutoff for the existence of a bond avoids the problem that arises when ideal gas atom end states are used. For typical values of the force constant for bond-stretching terms, i.e., $K > 100$ kcal/(mole Å^2)), the value of ϵ is $\epsilon \leq 1/100$. The TI integral would be best approximated if the simulation were run at the end points, i.e., at $\lambda = \epsilon$ or $1 - \epsilon$. However, this gives rise to two complications. At such small (large) values of λ, one may encounter van der Waals type end point problems for the nonbonded interaction terms.[36] In addition, the convergence behavior of $\langle \partial U_{\rm bond}/\partial\lambda\rangle_\lambda$ at e.g., $\lambda = 0.01$ or 0.99 is slow (S. Boresch & M. Karplus, unpublished results, refs 25 and 26). Both difficulties can be avoided since the analytical behavior of $\langle \partial U_{\rm bond}/\partial\lambda\rangle_\lambda$ is known (see eq 36). This is illustrated in Figure 3, where $\langle \partial U/\partial\lambda\rangle_\lambda$, as well as $\int_0^\lambda \langle \partial U/\partial\lambda\rangle_\lambda \, d\lambda$ are depicted as a function of λ. The full line is the analytical function. Data points taken from simulations of a one-dimensional harmonic oscillator with a force constant of 500 kcal/(mole Å^2) for various values of λ are overlaid in Figure 3a; in the case of a single one-dimensional harmonic oscillator the simulation is clearly well behaved, but this is not likely to be true in general. The plots show the strongly varying behavior of the integrand which will cause problems for numerical integration (as mentioned above) unless a large number of datapoints is used in the vicinity of the endpoint. In analogy to the approach proposed by Simonson[36] for the van der Waals endpoint problem, it suffices to calculate a few points at intermediate values of λ (say between 0.1 and 0.9) to fit these values to $C_o + C_1/\lambda$ or $C_o + C_1/(1 - \lambda)$ (where C_o and C_1 are constants) and to integrate this function analytically between ϵ and 1 (or 0 and $1 - \epsilon$). In principle, this type of procedure has to be repeated for each bonded term (bond, bond angle, and Urey−Bradley term) that is changed in the mutation, as the limiting value of $\lambda = \epsilon$ is a function of the respective force constant. Whether this is necessary to obtain accurate results if double free energy cycles are of interest has to be investigated.

In a number of applications of the BLOCK module of CHARMM[6,7] sizable bond free energy components were reported. As already mentioned, BLOCK is a dual topology method, and in its standard implementation bonded energy terms are scaled by λ. In the applications, no calculations were made at or close to the end points $\lambda = 0$ and 1. This avoided the convergence problems encountered in refs 25 and 26, but the validity of the results is not clear since the limits for the integral do not correspond to those that have been shown here to yield the correct values for bond making or bond breaking. The use of double differences with the same cutoff for λ should result in a cancelation of the errors in the two separate calculations. This is exact for an isolated harmonic oscillator and is expected to be approximately correct for the more complex cases considered. The effect of coupling with nonbonded terms can give nonzero contributions, but the results of calculations presented in the companion paper suggest that they are small. Thus, it is most likely that the significant bonded contributions obtained in refs 6 and 7 are due to noise in the calculations. Sun et al.[29] have commented on the calculations in ref 6. Repeating the computation of the free energy difference of unfolding $\Delta\Delta A_{\rm unf}$, between the Ala 96 mutant of barnase and wild type with a single topology methodology, they found only negligible contributions from bonded energy terms to $\Delta\Delta A_{\rm unf}$ from a component analysis. The difference between a single and a dual topology approach was identified as the main reason for the different components obtained, and the nature of the practical problems encountered when attempting to form or remove a bond were recognized on the basis of theoretical considerations and model calculations. While we agree with much of the analysis of Sun et al., three points need to be clarified. (i) Sun et al. used a single topology in a thought experiment to mimic dual topology (as already described in section 2b). While sufficient to identify the source of the problems resulting in the artificially large bond components of ref 6, this hybrid approach is not identical to the dual topology methods implemented in CHARMM. (ii) Sun et al. imply that the divergence of the integrand (eq 36) "may be removed when

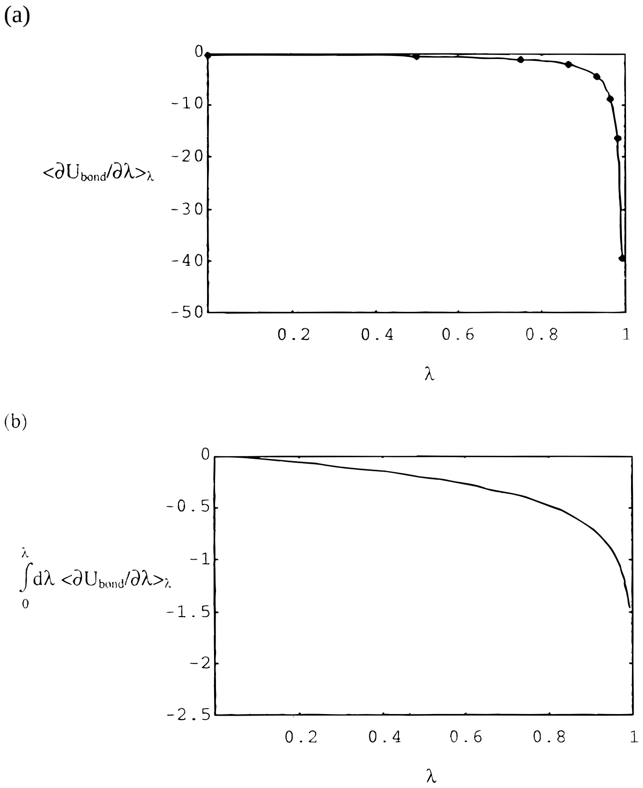

Figure 3. Theoretical behavior of $\langle \partial U_{\rm bond}/\partial\lambda\rangle_\lambda$ (Figure 3a) and $\int_0^\lambda \langle \partial U_{\rm bond}/\partial\lambda\rangle_\lambda \, d\lambda$ (Figure 3b) as a function of the coupling parameter λ for a harmonic oscillator with force constant $K = 500$ kcal/(mol Å^2) and equilibrium bond length $x_o = 1.526$ Å based on eq 36. The points in Figure 3a were obtained from simulation. Note that the plot ranges only between 0 and $1 - \epsilon$, where ϵ is given by eq 40.

nonlinear coupling is introduced between this vibrational mode and other modes of the system". Contrary to the van der Waals end point problem,[36] this is not the case here. The integral over eq A5 of ref 29 with $k_1 = 0$ (which is essentially equivalent to eq 36) is logarithmically divergent, regardless of the dependence on the coupling parameter. The correct solution, as shown in this chapter, lies in the use of physically appropriate limits of integration, i.e., ϵ and $1 - \epsilon$ rather than 0 and 1). Finally, (iii) while the different treatment of bonded terms in single and dual topology is emphasized, it is not made clear (as shown in section 2b and discussed further in section 2e) that the use of an ideal gas molecule end state is correct and does not omit any contributions to $\Delta\Delta A_{\text{unf}}$.

The analysis of bond stretching terms carries over to bond angle terms. Apparent theoretical problems can be removed by considerations similar to those used to obtain eq 39, i.e., by introducing appropriate limits of integration. From eq 34 integrating over a bond angle degree of freedom with zero force constant does not give the full volume V of the system, but the volume element expected for the Jacobian of the transformation from Cartesian into internal coordinates, i.e., $\int_0^{+\pi} \sin\theta \, d\theta = 2$. In practical applications, similar problems as for bond terms are observed. For $\lambda \rightarrow 0$ (or 1) the system will sample the full range of angles between $-\pi$ and $+\pi$; however, as most of these angle values correspond to high energies in the original system, $\langle \partial U_{\text{angle}}/\partial\lambda \rangle_\lambda$ becomes increasingly noisy and converges slowly (S. Boresch & M. Karplus, unpublished results). These difficulties can be avoided either by the use of an ideal gas molecule end state, i.e., angle terms are not scaled by λ, or by the introduction of a cutoff for the existence of the angle. The ϵ defined in eq 40 can also be used for bond angle terms. In free energy simulations using the CHARMM all-atom force field[74] an alchemical mutation may also change a bond angle degree of freedom described by a bond angle plus a Urey–Bradley term into one that is only described by a bond angle term (or vice versa). Since this is equivalent to a change in the effective bond angle force constant $K_{\theta,\text{eff}}$ rather than a removal of a bond angle degree of freedom (cf. the discussion given at the end of section 2c and in the Appendix), no special practical problems are encountered in this case.

2e. Meaning of Components in the Free Energy Differences of Solvation. Since the free energy change in TI is projected on the λ-dependent terms, care has to be used in interpreting the results, including that for the bonded contributions to the free energy of solvation. This is all the more important as the λ-dependent terms in the hybrid potential energy function vary not only with the path used for the simulations[8,19,32,33,51] but also with the simulation methodology (single or dual topology), cf. section 2b. To analyze the meaning of the bonded and the self-terms (see Introduction), we compare the chemical and alchemical paths in the thermodynamic cycle used to compute the free energy difference of solvation between ethane and methanol (Figure 1) with respect to the role of bond and bond angle free energy components. Numerical results for this model system are presented in the companion paper.[42]

Vertical (chemical paths ΔA_1 and ΔA_2 in Figure 1) and horizontal arrows (alchemical paths ΔA_3 and ΔA_4 in Figure 1) correspond to different processes. Along the chemical paths 1 and 2, the two solutes (e.g., ethane and methanol) are transferred from the gas phase into solution, so the hybrid potential energy function is of the form

$$U(\lambda) = U_{\text{solute}} + \Delta U_{\text{solute-water}}(\lambda) \tag{42}$$

and the free energy difference is equal to

$$\Delta A = \int_0^1 d\lambda \langle \partial \Delta U_{\text{solute-water}}/\partial\lambda \rangle_\lambda \tag{43}$$

since U_{solute} is independent of λ. There is no difference between a single and a dual topology calculation, as a single solute is transferred from the gas phase into solution in both cases. Along the chemical paths, the λ-dependent perturbation term in the potential energy function consists only of solute–solvent interactions; i.e., *only* nonbonded interactions (van der Waals and electrostatic terms in the usual force fields) are changed; consequently, *only* nonbonded free energy components are obtained for the free energy differences ΔA_1 and ΔA_2. It follows that the same is true for the free energy difference of solvation calculated as $\Delta\Delta A_{\text{solv}} = \Delta A_2 - \Delta A_1$.

The interaction of the solute with the solvent can alter the average bond lengths and bond angles and/or vibrational frequencies of the solutes. If any of these properties are different in the gas phase and in solution, they contribute to the free energy difference. Such contributions are projected onto the solute–solvent interaction terms in the chemical transformations. They are included in eq 43 through the Boltzmann weighting factor that depends on the full potential energy function, $U(\lambda)$ (see eqs 7 and 11). Thus, neither explicit bond(ed) free energy components nor, for that matter, components that have their origin in changes of intramolecular nonbonded interactions caused by structural changes due to solvation are obtained along the chemical paths. This corresponds to the analysis of Ben-Naim (e.g., ref 76) who argues that the influence of solvent on bond and bond angle terms (hard internal degrees of freedom) can be expressed as an additive term; i.e., the free energy of a solvated molecule is written as the free energy of the molecule in the gas phase plus the solvation contribution. Although an approximation, in general, such an approach is correct for systems described by the classical mechanical and pairwise additive potentials commonly used in computer simulations.[17,20,21]

The hybrid potential energy function describing the alchemical paths (the horizontal arrows in Figure 1) has the form

$$U(\mu) = U_{\text{o}} + \Delta U_{\text{solute}}(\mu) + \Delta U'_{\text{solute-water}}(\mu) \tag{44}$$

where we use μ instead of λ to indicate that the coupling parameter corresponds to a different process than that considered in eq 42, (i.e., the process described by eq 44 is that of transmuting ethane to methanol. The quantity $\Delta U_{\text{solute}}(\mu)$_corresponds to the change in intrasolute interactions and $\Delta U'_{\text{solute-water}}(\mu)$ gives the corresponding change in solute–solvent interactions. The $\Delta U'_{\text{solute-water}}(\mu)$ in eq 44 is not the same as $\Delta U_{\text{solute-solvent}}(\lambda)$ of eq 42; the former describes the change in solute–solvent interactions due to the alchemical mutation, whereas the latter describes the process of transferring ethane *or* methanol from the gas phase into solution. The alchemical free energy difference is given by

$$\Delta A = \int_0^1 \{ \langle \partial \Delta U_{\text{solute}}/\partial\mu \rangle_\mu + \langle \partial \Delta U'_{\text{solute-water}}/\partial\mu \rangle_\mu \} \tag{45}$$

A comparison of eqs 43 and 45 makes clear that different free energy components are to be expected. Equation 45 contains components that arise from the intramolecular changes in the solute which are not present in eq 43. In the ethane to methanol case, as well as most other practical applications, these include changes in bonded energy terms. A difference in the free energy of solvation calculated along the alchemical paths

ΔA_4 and ΔA_3 can contain contributions from the change in intrasolute interactions. If they are identical in solution and in the gas phase, they cancel in $\Delta\Delta A_{\mathrm{solv}}$. In a number of calculations mentioned in the Introduction it was, in fact, assumed that they were identical without verification.

The difference in free energy components along the alchemical and chemical paths reflects what has been called the *global* path dependence of thermodynamic integration.[32] The total double free energy difference of solvation is the same for the two paths. However, the explicit calculation of the free energy contributions from the changes in the intramolecular potential of the solute for the gas phase and solution along the alchemical paths makes it possible to determine the influence of solvent on these degrees of freedom (the self-terms), which is "hidden" along the chemical paths. In the ethane/methanol transmutation (Figure 1) the self-terms (or *intraperturbed-group interactions*[22]) contribute to $\Delta\Delta A_{\mathrm{solv}}$ only if they are not equal in ΔA_3 and ΔA_4. A self-term contribution to $\Delta\Delta A_{\mathrm{solv}}$ found along the alchemical paths provides information on the influence of solvent on the properties of the solutes. Such a separation of the intrasolute from the solute–solvent terms is not directly available from calculations along the chemical paths.

For the interpretation of components obtained along the alchemical paths ΔA_3 and ΔA_4, the differences between single and dual topology methods (the latter with and without scaling of bonded terms) have to be taken into account as well (section 2b). There are no bonded free energy components if an ideal gas molecule end state is used in dual topology simulations. It is important to note that calculations carried out in this manner *omit* vibrational and Jacobian factor free energy contributions, while pmf-type contributions are projected on the nonbonded free energy components. Thus, different single free energy differences for the gas phase (ΔA_3) and solution calculation (ΔA_4) are obtained, reflecting the different end states used in the simulations compared to a single topology or a dual topology calculation in which an ideal gas atom end state is used. However, as analogous contributions are omitted in the gas phase and in solution, the double free energy difference $\Delta\Delta A_{\mathrm{solv}}$ is not affected. This is the case even if coupling between bonded terms and nonbonded solute–solvent interactions is important. Similar to the chemical paths, any influence of the nonbonded terms on the intramolecular degrees of freedom is implicitly included in the Boltzmann factor that determines the distribution which contributes to the configuration integral (eqs 5, 7, and 11). Dual topology simulations in which bonded terms are scaled to a limiting value of λ as described in section 2d (ideal gas atom end state) are expected to include all bonded free energy contributions and, thus, to yield the same free energy differences in the gas phase and in solution that are obtained in single topology simulations.

3. Summary

Bond stretching and bond angle bending terms in the energy functions have led to problems in free energy simulations. The theoretical investigation of these terms reported here makes clear that several factors are essential in order to calculate them correctly and to obtain a meaningful interpretation of the results. The role and importance of bonded energy terms was shown to depend on the simulation methodology (single or dual topology). If a bond (angle) term is changed in an alchemical mutation using single topology, the resulting free energy difference arises from the three physical effects: (i) *vibrational* contributions from changes in the force constants, (ii) *Jacobian factor* contributions from changes in the equilibrium bond lengths and bond angles in the absence of nonbonded interactions, and (iii) *pmf-type* contributions which result from the change in nonbonded interaction (e.g., solute–solvent interactions) if the equilibrium (average) geometry of a molecule is altered. Applying constraints to keep the bonds rigid is a special case, in which vibrational terms are eliminated, but there can still be pmf-type and Jacobian factor free energy contributions if the equilibrium bond length of a constrained bond term (i.e., the target distance of the constraint) is changed. The same three contributions are obtained in dual topology methods when an ideal gas atom end state is used; however, the pmf-type free energy contribution is projected on nonbonded free energy components. In dual topology methods using an ideal gas molecule reference state, vibrational and Jacobian factor contributions are not included in the computed free energy difference. This results in different single free energy differences for individual steps of a thermodynamic cycle; however, identical double free energy differences are obtained.

In principle, all three contributions (vibrational, pmf-type, Jacobian factor) can couple with each other. Of particular interest is coupling between vibrational and pmf-type contributions. We reiterate that all three methodologies (single topology, dual topology with an ideal gas atom or molecule end state) include such effects correctly. Since Jacobian factor and vibrational contributions (but not coupling contributions!) are omitted in dual topology simulations using an ideal gas molecule reference state, a comparison of free energy components obtained with these three different methods allows one to distinguish these contributions. Further, vibrational degrees of freedom are absent in systems where bond terms are constrained, so simulations with and without constraints can be used to explore the importance of such coupling.

The origin of the problems arising if bonds or angles are broken or made was analyzed. Such problems can be avoided by use of a hybrid potential energy function in alchemical free energy calculations that does not scale the bonding terms. This can be achieved by the introduction of unphysical systems in the representation of the initial and/or final state. In single topology simulations, dummy atoms are used to keep the number of atoms in the system constant (cf. the situation in the ethane to methanol example, Figure 2a), and in all published work the force constants of the bond and bond angle terms connecting them to the rest of the system have been included. In dual topology simulations using an ideal gas molecule end state, the molecular fragment remains bonded harmonically to the rest of the system; all other interactions with the rest of the system are switched off. Thus, one obtains a "dummy molecular fragment" as the end state (cf. Figure 2b). Consequently, in both methodologies one or both endstates are not identical to the physical system they attempt to mimic. The unphysical modifications that may be necessary to accomplish an alchemical mutation have to be taken into account, which has not been done in all published work.[13,23] It was shown that if treated properly the artifacts of the simulation (i.e., the parts of the hybrid system that are not present in the real system they mimic) do not affect the result of the double free energy differences of interest. Alternatively, a limiting value of the coupling parameter (and thus the force constant) at which a bond or angle is considered broken can be used in the free energy simulation.

The focus of this paper has been the computation of correct double free energy differences defined by a thermodynamic cycle. In many cases this is the physically relevant quantity. The considerations and techniques presented here are equally useful when an absolute; i.e., single, free energy difference is required. In this case, even more attention has to be paid to the artifacts that may be introduced by the simulation methodology. As just discussed, bond and bond angle terms to dummy atoms are maintained in single topology calculations (e.g., refs 22, 24, and 54). If the contributions resulting from these terms are treated consistently, they cancel from a thermodynamic cycle. To obtain correct single free energy differences, the necessary correction can be calculated analytically[61] by use of the techniques developed by Herschbach et al.[66] In dual topology simulations the bond and bond angle terms of the two molecular fragments that represent the initial (reactant) and final (product) state, respectively, can be kept intact (ideal gas molecule end state) or they can be reduced to an appropriate threshold (ideal gas atom end state); details of the approach are discussed in section 2d. This resembles the single topology case, i.e., vibrational and Jacobian factor contributions (there are no others!) of the extraneous atoms bonded harmonically to the physical system (the "dummy molecular fragment") can be computed analytically. The latter approach gives essentially an absolute free energy difference. The single contribution omitted by the method outlined in section 2d is a correction for the unit volume implicit in ϵ (eq 40 instead of the true volume of the (simulation) system. The need for such a correction is reminiscent to the need to choose a standard state in the calculation of absolute binding affinities.[77]

As discussed in the Introduction, there is considerable confusion in the literature regarding both theoretical and practical aspects of the role of bonded terms in free energy simulations. We defer a detailed discussion of this point to the end of the companion paper, in which results of model calculations illustrating the conclusions drawn here are reported.

Acknowledgment. We thank Tom Simonson, Georgios Archontis and Arnaud Blondel for helpful discussions. This work was supported by a grant from the National Institute of Health.

Appendix: Analytical Calculation of a Free Energy Difference Involving the Simultaneous Change of a Bond Angle and Urey−Bradley Term

To illustrate the difficulty arising if a bond angle degree of freedom is determined simultaneously by an angle term in θ and a Urey−Bradley term in the 1−3 distance, we consider again the nonlinear triatomic molecule used at the end of section 2c. The four internal coordinates r, r′, r'', and θ are not independent since there are only 3 degrees of freedom. The (gas phase) free energy of a polyatomic molecule within the rigid rotor, harmonic oscillator (RRHO) approximation is given by[61,66]

$$A = -k_B T \ln \frac{(2\pi k_B T)^{3N/2-3}}{|F_S|^{1/2}} \prod_i J_{iS} \qquad \text{(A1)}$$

The J_{iS} are the Jacobian factors discussed in the main text, and $|F_S|$ is the determinant of the force matrix in the internal coordinates used, i.e., it has elements $F_{ij,S} = \partial^2 U/\partial r_i \partial r_j$, where U is the potential energy of the system and r_i and r_j are internal coordinates. As long as the number of internal coordinates and the number of degrees of freedom are the same, the force matrix is diagonal (e.g., for a triatomic system described by two harmonic bond stretching terms and one bond angle energy term, $|F_S| = KK'K_\theta$). To be able to use eq A1, the internal coordinates have to be nonredundant. This is not the case if the Urey−Bradley energy term is added since $r'' = r''(r,r',\theta)$. Once one decides on a subset of nonredundent coordinates, the J_{iS} can be used as before; however, there are off-diagonal elements in the force matrix since the potential energy is expressed as a function of all four (redundant) coordinates. For the triatomic model system, evaluating the elements of F_S is tedious, but straightforward: One first carries out the double differentiation, then, in accord with the RRHO approximation, simplifies the resulting expressions by replacing r'' by r_o'', i.e., the equilibrium 1−3 distance. A program capable of symbolic manipulation aids in the computations.

TABLE A1

parameters	initial state	final state
$r_{o,23}$(Å), K_{23}(kcal/(mol Å^2))	1.5/250.0	1.0/500.0
θ_o(deg.), K_θ(kcal/(mol rad^2))	110./25.0	110./50.0
$r_{o,13}$(Å), K_{UB}(kcal/(mol Å^2))	2.88/25.0	2.52/25.0

Extending this exact approach to larger systems quickly becomes very complicated. The following approximations may, therefore, be useful for estimates in larger systems. As Urey−Bradley terms in the CHARMM force field[74] are an additive correction to the standard bond angle terms, we replace the bond angle force constant K_θ by an effective bond angle force constant $K_{\theta,\text{eff}}$ which includes the effect of the Urey−Bradley term in the potential energy function in an approximate manner, i.e., $U_{\text{angle}} + U_{\text{UB}} \approx K_{\theta,\text{eff}}(\theta - \theta_o)^2$ (cf. eq 35 in section 2c). This maintains the diagonal form of F_S. The simplest possible expression for $K_{\theta,\text{eff}}$ uses the sum of bond angle and Urey−Bradley force constant, i.e., one writes $K_{\theta,\text{eff}} = K_\theta + K_{\text{UB}}$. A much better approximation can be obtained from a comparison of the diagonal force matrix in the absence of the Urey−Bradley term and the full expression with off-diagonal elements with the Urey−Bradley term. For typical values of the force constants, the largest change can be expected for $\partial^2 U/\partial^2\theta$, i.e., the diagonal element for the bond angle degree of freedom. It is given by

$$F_{\theta\theta} = \partial^2 U/\partial^2\theta = K_\theta + K''({r_o}^2 {r_o'}^2 \sin^2\theta/{r_o''}^2) \approx K_{\theta,\text{eff}} \qquad \text{(A2)}$$

and provides an expression for $K_{\theta,\text{eff}}$ that takes into account the interdependence of the four internal coordinates. The procedure just outlined blurs the distinction between vibrational and Jacobian factor contributions. However, this only reflects the dependence of the 1−3 distance and the bond angle θ on each other. An approximate separation can be accomplished by comparing the results obtained with $K_{\theta,\text{eff}} = K_\theta + K_{\text{UB}}$ to those with $K_{\theta,\text{eff}}$ on the basis of eq A2, since the effect which the change of equilibrium geometry (r_o, r_o', θ_o) has on the free energy contribution from a Urey−Bradley term is included only in the latter.

We conclude this appendix with a numerical example for a model triatomic molecule. Table A1 lists the bonded parameters that are different in the initial and final state; there are no nonbonded interactions in the system and the parameters for the 1−2 bond are K_{12}=250 kcal/(mole Å^2), $r_{o,12}$=2 Å.

Thus, aside from the change in the 2−3 bond stretching term, which is straightforward to calculate, we expect a vibrational contribution from the bond angle term, since K_θ changes, and a Jacobian factor contribution from the Urey−Bradley term because of the change in 1−3 distance.

Approaches for computing analytically the free energy difference for such a system are described in ref 61. The reference result for this alchemical mutation is obtained by adding the contribution due to the change in moment of inertia ΔA_I, which is not negligible for such a small system, to the

TABLE A2

ΔA_{I}	0.315
ΔA_{NM}	0.459
$\Delta A_{exa} = \Delta A_{I} + \Delta A_{NM}$	0.774
ΔA_{NUB}	0.896
ΔA_{M1}	0.810
ΔA_{M2}	0.769
ΔA_{M3}	0.774

free energy difference obtained from a normal-mode analysis of the initial and final state, ΔA_{NM}; see eq 14 of ref 61. (This approach assumes that the masses are not changed.) The alternative approach of adding the vibrational and Jacobian factor free energy differences (cf. Equation 31) has to take into account the redundancy of the internal coordinates as described above. We compare the reference result (ΔA_{exa}) and the equally exact result based on eq A1 using the full force matrix with all off-diagonal elements (labeled M3) with those obtained ignoring the Urey–Bradley term (ΔA_{NUB}) and the two approximations including the Urey–Bradley term. The simpler approach assuming that $K_{\theta,eff} = K_{\theta} + K_{UB}$ is labeled M1, whereas the second, more elaborate, method (M2) uses expression eq A2 for $K_{\theta,eff}$. The results obtained are summarized in Table A2 (all values are in kcal/mol).

In Table A2, we first list separately the two contributions (ΔA_{I}, ΔA_{NM}) to the reference result ΔA_{exa}. The fourth entry ΔA_{NUB} omits the Urey–Bradley term and was obtained from the Jacobian contribution due to the change in the 2–3 bond length plus the vibrational contributions due to the changes in K_{23} and K_{θ}, cf. Table A1. The resulting free energy difference is 15.7% too high. Although a crude approximation (K_{θ} is simply replaced by the sum $K_{\theta} + K_{UB}$), the agreement of ΔA_{M1} with ΔA_{ref} is much better (4.6% too high). The second approximate free energy difference (ΔA_{M2}) is accurate within 1%. The results of the two exact methods (ΔA_{exa}, ΔA_{M3}) agree as they must. Finally, it is interesting to compare the values for $K_{\theta,eff}$ in M1 and M2 to those of K_{θ} (Table A1). Using method M1, $K_{\theta,eff}$ equals 50 and 75 kcal/(mol rad^2) for initial and final state, respectively. When using approach M2, these values change to 48.9 and 63.9 kcal/(mol rad^2), reflecting the influence of the changed 2–3 bond length that is omitted in method M1.

References and Notes

(1) Brooks, C. L., III; Karplus, M.; Pettitt *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; *Advances in Chemical Physics* LXXI; John Wiley & Sons: New York, 1988.

(2) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431–492.

(3) McCammon, J. A.; Straatsma, T. P. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407.

(4) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395–2417.

(5) Dang, L. X.; Merz, K. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 8505–8508.

(6) Prevost, M.; Wodak, S. J.; Tidor, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10880–10884.

(7) Tidor, B.; Karplus, M. *Biochemistry* **1991**, *30*, 3217–3228.

(8) Lau, F. T. K.; Karplus, M. *J. Mol. Biol.* **1994**, 1049–1067.

(9) Miyamoto, S.; Kollman, P. A. *Proteins* **1993**, *16*, 226–245.

(10) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. *Science* **1989**, *244*, 1069–1072.

(11) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050–3054.

(12) Tobias, D. J.; Brooks, C. L., III *J. Phys. Chem.* **1992**, *96*, 3864–3870.

(13) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564–568.

(14) Elcock, A. H.; Richards, W. G. *J. Am. Chem. Soc.* **1993**, *115*, 7930–7931.

(15) Elofsson, A.; Kulinski, T.; Rigler, R.; Nilsson, L. *Proteins* **1993**, *17*, 161–175.

(16) Sun, Y.; Spellmeyer, D.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1992**, *114*, 6798–6801.

(17) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(18) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281–283.

(19) Boresch, S.; Archontis, G.; Karplus, M. *Proteins* **1994**, *20*, 25–33.

(20) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(21) Jorgensen, W. L.; Swenson, C. *J. Am. Chem. Soc.* **1985**, *106*, 765–784.

(22) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532–4546.

(23) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. *Nature* **1987**, *328*, 551.

(24) Van Gunsteren, W. F. In *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*; Van Gunsteren, W. F., Weiner, P. K., Eds.; Escom Science: Netherlands, 1989; pp 27–59.

(25) Elofsson, A.; Nilsson, L. *Mol. Simul.* **1993**, *10*, 255–276.

(26) Eriksson, M. A. L.; Nilsson, L. *J. Mol. Biol.* **1995**, *253*, 453–472.

(27) Harris, D.; Loew, G. *J. Comput. Chem.* **1996**, *17*, 273–288.

(28) Rao, B. G.; Kim, E. E.; Murcko, M. A. *J. Comput. Aid. Mol. Des.* **1996**, *10*, 23–30.

(29) Sun, Y.-C.; Veenstra, D. L.; Kollman, P. A. *Protein Eng.* **1996**, *9*, 273–281.

(30) Mark, A. E.; van Gunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167–176.

(31) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 577–585.

(32) Boresch, S.; Karplus, M., *J. Mol. Biol.* **1995**, *254*, 801–807.

(33) Archontis, G.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 11246–11260.

(34) Shi, Y.; Mark, A. E.; Wang, C.; Huang, F.; Berendsen, H. J. C.; van Gunsteren, W. F. *Protein Eng.* **1993**, *6*, 289–295.

(35) Dill, K. A. *J. Biol. Chem.* **1997**, *272*, 701–704.

(36) Simonson, T. *Mol. Phys.* **1993**, *80*, 441–447.

(37) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529–539.

(38) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025.

(39) Simonson, T.; Archontis, G.; Karplus, M. *J. Phys. Chem.* **1997**, *41*, 8347–8360.

(40) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem.* **1996**, *100*, 1206–1215.

(41) Archontis, G.; Simonson, T.; Moras, D.; Karplus, M. *J. Mol. Biol.* **1998**, *275*, 823–846.

(42) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1988**, *103*, 119.

(43) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.

(44) Hansen, J.-P.; McDonald, I. R. *Theory of Simple Liquids*; Academic Press: London, 1986.

(45) Zwanzig, R. W.; *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(46) McQuarrie, D. A. *Statistical Mechanics*; Harper & Row: New York, 1976.

(47) Straatsma, T. P.; Zacharias, M.; McCammon, J. A. In *Computer Simulations in Biological Systems*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Escom Science: Netherlands, 1994.

(48) Van Gunsteren, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. In *Computer Simulations in Biological Systems*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Escom Science: Netherlands, 1994.

(49) Simonson, T.; Brünger, A. T. *Biochemistry* **1992**, *31*, 8661–8674.

(50) Hodel, A.; Simonson, T.; Fox, R. O.; Brünger, A. T. *J. Phys. Chem.* **1993**, *97*, 3409–3417.

(51) Brady, G. P.; Sharp, K. A. *J. Mol. Biol.* **1995**, *254*, 77–85.

(52) Pearlman, D. A. *J. Phys. Chem.* **1994**, *98*, 1487–1493.

(53) Cross, A. J. *Chem. Phys. Lett.* **1986**, *128*, 198–202.

(54) Hermans, J.; Yun, R. H.; Anderson, A. G. *J. Comput. Chem.* **1992**, *13*, 429–442.

(55) Resat, H.; Mezei, M. *J. Chem. Phys.* **1993**, *99*, 6052–6061.

(56) Jorgensen, W. L. *BOSS*, version 3.2, Department of Chemistry: Yale University, New Haven, 1992.

(57) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. *AMBER 4.1*; University of California: San Francisco, 1995.

(58) Van Gunsteren, W. F.; Berendsen, H. J. C. *GROningen MOlecular Simulation (GROMOS) Library Manual*; Laboratory of Physical Chemistry, University of Groningen: Groningen, 1987.

(59) Straatsma, T. P.; Zacharias, M.; McCammon, J. A. *Chem. Phys. Lett.* **1992**, *196*, 297–302.

(60) Wang, L.; Hermans, J. *J. Chem. Phys.* **1994**, *100*, 9129–9139.

(61) Boresch, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 5145−5154.

(62) Miranker, A.; Karplus, M *Proteins* **1995,** *23*, 472−490.

(63) Verkhivker, G.; Elber, R.; Nowak, W. *J. Chem. Phys.* **1992**, *97*, 7838−7841.

(64) Bash, P. A.; Field, M. J.; Karplus M. *J. Am. Chem. Soc.* **1987**, *109*, 8092−8094.

(65) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Verlag Chemie Publishers: New York, 1996; Vol. 7, Chapter 3, pp 119−185.

(66) Herschbach, D. R.; Johnston, H. S.; Rapp, D. *J. Chem. Phys.* **1959,** *31*, 1652.

(67) Severance. D. L.; Essex, J. W.; Jorgensen, W. L. *J. Comput. Chem.* **1995**, *16*, 311−327.

(68) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(69) Karplus, M.; Kushick J. N. *Macromolecules* **1981**, *14*, 325−332.

(70) Pearlman, D. A. *J. Chem. Phys.* **1993**, *98*, 8946−8957.

(71) Guo, H.; Karplus, M. *J. Phys. Chem.* **1992,** *96*, 7273.

(72) de Souza, L. E. S.; Guerin, C. B. E.; Ben-Amotz, D.; Szleifer, I. *J. Chem. Phys.* **1993**, *99*, 9954−9961.

(73) Pettitt, B. M.; Karplus, M. *J. Am. Chem. Soc.* **1985**, *107*, 1166−1173.

(74) MacKerell, A. D. Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L. Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E. III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Workiewicz-Kuczera, J.; Yinb, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(75) Van Gunsteren, W. F. *Mol. Phys.* **1980**, *40*, 1015.

(76) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016−2027.

(77) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.